UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

CRT-611

LOCATING POSTAL RELAY BOXES USING A SET COVERING ALGORITHM

by

Jocelyn Bouliane*
and
Gilbert Laporte**

* École Polytechnique de Montréal, C.P. 6079, Succursale A, Montréal (Québec) CANADA H3C 3A7.

** Centre de recherche sur les transports, Université de Montréal, C.P. 6128, Succursale A, Montréal (Québec) CANADA H3C 3J7.

Université de Montréal
Centre de recherche sur les transports - Publication #611
February 1989

Dépôt légal - 1er trimestre 1989

ABSTRACT

The problem of locating postal relay boxes is considered and modeled as a set covering algorithm. It is then solved by means of a simple heuristic solution procedure.

Key words : relay boxes, set covering, location.

- - - - - -

RÉSUMÉ

On considère le problème de la localisation des boîtes de relais dans un réseau postal. Le problème est modélisé comme un problème de recouvrement d'ensembles et résolu à l'aide d'un algorithme heuristique simple.

Mots-clés : boîtes de relais, recouvrement d'ensembles, localisation.

1. INTRODUCTION

Over the last decade, some members of our research centre have undertaken a number of studies in cooperation with Canada Post Corporation (CPC) in order to improve some of their operations. These include post-box collection routes[1] and door-to-door mail delivery[2]. The present study deals with the location of relay boxes along postmen's routes in Canadian cities. Relay boxes are made necessary by the fact that postmen can only carry a limited amount of mail at any time (at most 35 lbs) and must sometimes replenish their mail bags at specific points on their route. It is therefore important to determine relay box locations so that postmen never carry more than the prescribed weight of mail and do not have to walk unnecessarily when they run out of mail. Moreover, since these boxes must be visited each day by a delivery van before postmen start their routes, the number and location of relay boxes has a direct impact on the cost of this operation. In fact, post-box location also influences this cost since frequently the same vehicle is used to fill relay boxes and to empty post-boxes. From a more general standpoint, most postal operations are intertwined so that they should ideally be modelled and optimized simultaneously (several of these combined problems are location-routing problems[3]). This is of course more often than not impractical and the various subproblems are to a large extent solved independently. Our problem was solved by explicitly taking into account the weight constraint imposed on postmen, but it was also decided to minimize the number of relay box locations since by and large, if there are fewer relay box sites, it will be cheaper to replenish them.

2. THE MODEL

For the purpose of mail delivery, Canada is partitioned into territories (defined from postal codes) assigned to postal stations. Our problem

is solved at this level. Within a given territory, postmen's routes have already been determined (see Roy and Rousseau[2]) and mail volumes associated with street or road segments are estimated through statistical surveys. A territory is represented by an oriented graph $G=(N,A)$ where $N=\{n_j\}$ is the node set and $A=\{a_i\}$ is the arc set. Most often, nodes coincide with street or road intersections; they are potential locations for relay boxes. Note that the size of these relay boxes is not taken into account since several of them can be located side by side if the need arises. Arcs represent street or road segments. Let w_{ik} be the weight (in lbs) delivered by postman k on arc a_i and let w be the maximal weight (in lbs) which may be carried at any time by a postman (currently, w=35). Any given arc a_i can be traversed by several postmen but unless it does not require any mail delivery, it is serviced by only one of them. Therefore, for a given i, w_{ik} can take a positive value for at most one index k. Now consider postman k and his associated route including node n_j, arc a_{i_t} and the sequence of arcs a_{i_1}, a_{i_2}, ..., a_{i_t}, ..., where n_j is the origin of arc a_{i_1}. Arc a_{i_t} is said to be <u>covered</u> by node n_j if and only if it can feasibly be serviced by postman k replenishing at n_j, i.e. if $\sum_{s=1}^{t} w_{i_s k} \leq w$. Thus the problem can be formulated as a set covering problem. Let y_j be a binary variable equal to 1 if and only if a relay box is located at node n_j and δ_{ij}, a binary coefficient equal to 1 if and only if arc a_i is covered by node n_j. The problem is then

$$\text{(SCP)} \qquad \text{minimize} \qquad z = \sum_{n_j \in N} y_j \tag{1}$$

$$\text{subject to} \qquad \sum_{n_j \in N} \delta_{ij} y_j \geq 1 \qquad (a_i \in A) \tag{2}$$

$$y_j \in \{0,1\} \qquad (n_j \in N). \tag{3}$$

3. THE SET COVERING ALGORITHM

Rather than solving the set covering problem optimally, we decided to use the heuristic proposed by Salkin[4]. This decision was largely motivated by the size of the problem which can become quite large in practice (a real-life problem typically contains 150 nodes and 400 arcs). We also believe that it makes little sense to optimize too finely an objective which is largely a surrogate function and does not reflect an exact cost. Before describing the core of the algorithm, note that it is always possible to reduce the problem size.

3.1 Reduction Techniques [5]

If some constraint s contains only one non-zero coefficient δ_{sj}, fix y_j at 1 and remove all constraints r for which $\delta_{rj}=1$. Repeat this operation as often as possible. Now, if two constraints s and t are such that $\delta_{sj} \leq \delta_{tj}$ for all j, then constraint t can be dropped since it is satisfied whenever constraint s is. □

Note that in our problem, the effect of the first reduction technique is to locate a relay box at the beginning of every postman's route and to eliminate the arcs covered by this relay box. Hence, in this context, it is always possible to reduce the problem size as long as care is taken to distinguish the first and last node of cyclic routes.

Now consider SCP', the relaxed version of SCP obtained by dropping the integrality requirements, i.e. by using

$$0 \leq y_j \leq 1 \qquad (n_j \in N) \qquad (3')$$

instead of (3). (Note that the condition $y_j \leq 1$ can be relaxed since it is automatically satisfied at optimality[4].) The rounded-up solution value $\lceil \underline{z} \rceil$ to the relaxed problem is obviously a valid lower bound on the SCP optimum z^*. And if the solution $y=(y_1, y_2, \ldots, y_{|N|})$ to the relaxed problem is integer, it is also feasible and optimal for SCP. Otherwise, a feasible solution of value $\bar{z} \geq z^*$ is constructed from y. Again, this solution is optimal for SCP whenever $\bar{z} = \lceil \underline{z} \rceil$. The integer solution obtained by replacing by 1 every non-integer component of y is obviously feasible. But, as shown by Lemke et al.[6], a better integer solution can sometimes be obtained by successively applying the following two techniques.

3.2 Improved Integer Solution Technique

Extract from SCP the variables which are currently non-integer and the constraints which are explicitly satisfied by these variables. This extraction process yields a binary coefficient matrix E. Repeat the following operation until all rows of E have been eliminated: fix at 1 the (fractional) variable corresponding to the column of E having the largest number of 1's; remove from E all rows corresponding to the constraints which have become satisfied. When E has no more rows, fix at 0 all remaining fractional variables. □

3.3 Extreme Point Integer Solution Technique

Let now $y=(y_1, y_2, \ldots, y_{|N|})$ be an integer feasible solution to SCP. Define S as the set of row numbers corresponding to constraints which are satisfied as equalities and $E=(e_{sj})$, as the coefficient matrix associated with these constraints. Then, if there exists an index j such that $y_j=1$ and $e_{sj}=0$ for all $s \in S$, variable y_j can be set equal to 0. This process is repeated as many times as this condition can be satisfied. □

The algorithm used to solve SCP is then:

Step 1: Apply the reduction technique.

Step 2: Solve the relaxed problem SCP' by means of the simplex algorithm. If the solution (of value $\underline{z}$) is integer, it is then optimal; stop.

Step 3: Derive a feasible solution to SCP by means of the improved integer solution technique. If the solution value $\bar{z}$ is equal to $\lceil \underline{z} \rceil$, the optimum has been found; stop.

Step 4: Repeatedly apply the extreme point integer solution technique until no improvement is possible. Stop. □

4. COMPUTATIONAL RESULTS

The algorithm was coded in FORTRAN and included as a subroutine the SAS/OR Simplex code. The computer used was the École Polytechnique IBM 7171.

The solution method just decribed was first applied to the H2B territory of the Montreal "Henri-Bourassa" postal station. This problem contains 138 nodes and 384 arcs. The mail volumes (weights) on the arcs were taken as the maximal daily figures provided by CPC, increased by 10% to allow for further variations. The problem was solved in less than 12 seconds, after 429 simplex iterations and 31 successful applications of the extreme point integer solution technique. A feasible solution with 13 relay boxes and a lower bound $\underline{z}$ of 12 was found. This compares more than favourably with the 33 relay boxes currently used by CPC in the H2B territory.

Sensitivity analyses were then performed by multiplying the mail volumes by a factor α varying between 0.1 and 2.0. The results, reported in table 1, show that problems with a large value of α tend to be harder to

solve and, as expected, require more relay boxes. Out of the 20 problems which were attempted, a proven optimal solution was found 17 times.

Further problems of the same size (138 nodes and 384 arcs) were generated, with mail volumes in the same range as in the original H2B problem. The results, reported in table 2, show that these random problems are in general more difficult to solve than the H2B problem: computation times are higher and a proven optimum is now obtained 45% of the time. The reason for this phenomenon is probably that the graph structure of these random problems is more complex than that of real ones.

Table 1: Computational results for the H2B territory

α	No. of simplex iterations	Lower bound (value z of SCP' solution)	Best known feasible solution value (*: optimal)	Computation time (seconds)
0.1	337	9.0	9*	10.16
0.2	337	9.0	9*	10.17
0.3	337	9.0	9*	10.16
0.4	337	9.0	9*	10.16
0.5	337	9.0	9*	10.15
0.6	378	10.0	10*	10.50
0.7	371	10.0	10*	10.67
0.8	444	10.0	10*	11.48
0.9	464	10.0	10*	11.81
1.0	460	12.0	13	11.97
1.1	509	12.5	13*	12.81
1.2	501	14.0	15	13.08
1.3	512	16.0	17	13.38
1.4	522	16.8	17*	13.52
1.5	522	17.3	18*	13.44
1.6	582	19.0	19*	13.76
1.7	646	21.5	22*	15.00
1.8	536	22.1	23*	14.04
1.9	594	22.1	23*	15.39
2.0	563	22.6	23*	14.26

Table 2: Computational results for random problems

α	No. of simplex iterations	Lower bound (value z of SCP' solution)	Best known feasible solution value (*: optimal)	Computation time (seconds)
1	904	8.3	9*	25.90
2	724	9.0	10	19.34
3	859	7.4	10	25.83
4	792	9.3	11	24.60
5	883	8.0	8*	26.39
6	888	8.5	9*	25.67
7	815	10.0	11	21.97
8	834	9.3	10*	23.35
9	828	9.2	10*	23.65
10	662	9.5	11	18.80
11	814	10.3	12	22.45
12	660	10.0	10*	18.00
13	824	8.0	9	22.39
14	941	9.0	11	25.89
15	853	8.7	11	23.89
16	899	8.0	9	24.99
17	707	9.1	10*	20.11
18	845	7.3	9	23.99
19	877	8.0	8*	24.05
20	692	8.3	9*	21.76

5. CONCLUSION

We have shown that the relay box location problem encountered by CPC can be modeled as a set covering problem and solved quite satisfactorily by means of a simple (and quite old) solution technique described in text-books. The technique was applied to the H2B postal territory data and it was shown that a substantial reduction in the number of relay boxes could be achieved. The interest of this result lies not so much in the saving directly associated with the cost of relay boxes, but rather in the lower expected replenishment cost on a daily basis. A pilot project is now being undertaken by our group to combine into a single algorithm the location and routing aspects of this problem and to better assess the true savings.

ACKNOWLEDGEMENTS

This work was partially supported by F.C.A.R. grant 89EQ3084. Thanks are due to Mr. Philippe-Éric Landry of Canada Post Corporation for his assistance and to Dr. Michel Gendreau for his helpful advice.

REFERENCES

1 G. Laporte, S. Chapleau, P.-É. Landry and H. Mercure (1989) An algorithm for the design of mailbox collection routes in urban areas. Transportation Science B (forthcoming).

2 S. Roy and J.-M. Rousseau (1989) The capacitated Canadian postman problem. INFOR (forthcoming).

3 G. Laporte (1988) Location-routing problems. In Vehicle Routing: Methods and Studies (B.L. Golden and A.A. Assad, Eds.), North-Holland, Amsterdam, 163-198.

4 H.M. Salkin (1975) Integer Programming. Addison-Wesley, Reading, Massachusetts.

5 R. Garfinkel and G.L. Nemhauser (1972) Optimal set covering: a survey. In Perspectives on Optimization: A Collection of Expository Articles (A.M. Geoffrion, Ed.), Addison-Wesley, Reading, Massachusetts, 164-193.

6 C. Lemke, H. Salkin and K. Spielberg (1971) Set covering by single branch enumeration with linear programming subproblems. Operations Research 19, 998-1022.